



Open camera or QR reader and scan code to access this article and other resources online.



Mark Swerdlow, MS

Submitted for publication August 6, 2022.
Accepted in revised form January 31, 2023.

*Correspondence: Department of Surgery,
Keck School of Medicine of USC, 1520 San Pablo
Street, Suite 4300, Los Angeles, CA 90033, USA
(e-mail: mswerdlo@usc.edu).

Reliability of an AI-Powered Application Across Different Mobile Devices for Assessment of Chronic Wounds

Mark Swerdlow,* Jessica Lo, and David G. Armstrong

Department of Surgery, Keck School of Medicine of USC, Los Angeles, California, USA.

Objective: Evaluate the inter- and intrarater reliability of a wound assessment tool in iPhone 12 and 13 mini modalities against a validated iPad mini/Structure Sensor configuration.

Approach: We assessed a wound measurement application (eKare inSight®) for result consistency in patients presenting with wounds. Assessments were analyzed using a two-way analysis of variance. Intraclass correlation coefficient (ICC) was computed for intrarater ($ICC_{1,1}$) and inter-rater ($ICC_{2,1}$) analysis using a two-way random effects model. Paired *t*-test assessed the statistical difference between measurement methods.

Results: Forty-two lesions were analyzed with surface areas ranging from 0.2 to 23 cm² (average 4.33 ± 5.44 cm²). A high level of reliability was observed for repeat wound area measurements by the same examiner ($ICC_{1,1} = 0.997$) and between examiners with iPhone 13 mini ($ICC_{2,1} = 0.998$). There was no significant difference between iPhone 12 and iPad mini/Structure Sensor ($p = 0.78$) or between iPhone 13 mini and iPhone 12 ($p = 0.22$). Minimal difference existed between iPhone 13 mini and iPad mini/Structure Sensor ($p = 0.049$, Cohen's $d = 0.01$).

Innovation: Increased pervasiveness of smartphones in clinical care, coupled with advances in smartphone imaging and machine learning, allows for a potential solution to the problem of fast and accurate wound measurements. The application investigated produces wound measurement results quickly and with demonstrated accuracy. It does not require a calibration sticker or reference marker and allows for automatic wound boundary delineation.

Conclusion: The results of this study suggest that a digital planimetry mobile application may offer high levels of reliability across devices and users.

Keywords: imaging, acute wound, chronic wound, digital planimetry

INTRODUCTION

ACUTE AND CHRONIC WOUNDS IMPOSE a substantial burden on patients, caregivers, and health care systems, affecting >8 million Americans and costing >28.1 billion annually.¹ It is estimated that chronic wounds will affect 1–2% of the population of developed countries at some point in their lifetime and up to a third of people with diabetes.^{2–4} Many

wounds have complex etiologies, with trauma, pressure injury, diabetes, and vascular insufficiency frequently playing contributory roles. Patients suffering from wounds may experience a decreased quality of life, pain, reduced mobility, life-threatening infection, disfigurement, and depression.^{5,6}

Accurate and efficient wound measurement remains clinically

challenging. It is not unusual for the same provider to use different devices across clinics or for wounds to not be measured due to time constraints. There is a need for accessible wound measuring tools that provide clinicians and patients with accurate wound characteristics to track wound healing and determine treatment efficacy. Recently, digital planimetry has advanced to incorporate numerous factors that affect wound measurement, including lighting, distance, angle, and curvature. Digital imaging tools have proven to be cost-effective, accessible, noninvasive, and time-efficient approaches to wound measurement.⁷ With the ubiquity of smartphones and tablets, digital imaging that includes two-dimensional and three-dimensional (3D) analysis could be available to health care providers globally. However, these tools must be validated against a gold standard and compared across different models.

CLINICAL PROBLEM ADDRESSED

Traditional methods of wound measurement are often time-consuming or inaccurate.

Wound tracing sheets are reliable but time-intensive and require contact with the wound, which increases chances of infection, wound bed damage, warping of wound shape, and patient discomfort.⁸ Calculating wound characteristics from in-image rulers is efficient but inaccurate. Quantifiable measurement data help to inform clinical decision-making and provide patients with tangible evidence of healing progress. The increasing use of mobile phones in clinical care coupled with advances in smartphone imaging and machine learning allows for a potential solution to this problem.

MATERIALS AND METHODS

This clinical validation study evaluated the inter- and intrarater reliability of a phone-based 3D digital imaging solution (inSight®; eKare Inc., Fairfax, VA) against a previously validated inSight configuration with iPad mini/Structure Sensor (inSight; eKare Inc.).^{9–16}

eKare App

The inSight CR app is an FDA-registered and CE-marked solution for assessment and documentation of lesions in a HIPAA- and GDPR-compliant environment. It uses computer vision and machine learning to measure wounds and create wound-bed topography renderings. Study images were stored in the app. An electronic laboratory notebook was not used. Previously, inSight has been validated in preclinical and clinical settings.^{8–14}

After signing in, the user creates a new patient entry and specifies wound location for the patient. Subsequently, they press the camera button to activate the back-facing camera and begin a new image entry for the patient. The user points the smartphone camera at the wound and clicks the standard smartphone circular image capture button. When the image appears on the next screen, the user uses one finger to circle around the wound and then swipe on the wound itself, which activates the automatic wound boundary detection algorithm. After pushing the “Measure” button, the measurement screen displays wound characteristics such as length, width, area, depth, and volume.

Patient recruitment and image collection

The study was approved by the University of Southern California (USC) Biomedical Institutional Review Board. Patients with active wounds, defined as a break in the epidermis, at least 0.2 cm² were recruited and consented from the USC Keck Physical Therapy Wound Care and Limb Preservation clinics (Los Angeles, CA). Patient wound images were collected during routine clinic visits during a data collection period of 40 days.

Study design

A total of five images per wound were captured by two users using three devices (Fig. 2). All images were captured using the inSight CR app. Examiner A assessed the wound with an iPad mini/Structure Sensor, an iPhone 12, and twice with an iPhone 13 mini. Examiner B assessed the wound once with an iPhone 13 mini. All images were captured within a 2-min interval. The iPad image was considered the reference image due to previously confirmed validity of the iPad mini/Structure Sensor combination. Both examiners were medical students with minimal prior experience using inSight and had no characteristic differences between them.

Data analysis

The test and retest assessments were performed on patients presenting with active wounds. Wound area measurements from each five-image series were calculated in the inSight CR app. A sufficient number of wound images were captured to have an adequate power to detect a test–retest correlation of at least 0.50. To maintain a family-wise type I error of 0.05, 28 wounds were needed to provide 80% power. After accounting for attrition at roughly 10%, we aimed to collect data for 31 wounds but instead collected data for 42 wounds (Fig. 1).

The primary outcome was inter- and intrarater reliability of a phone-based solution in assessment

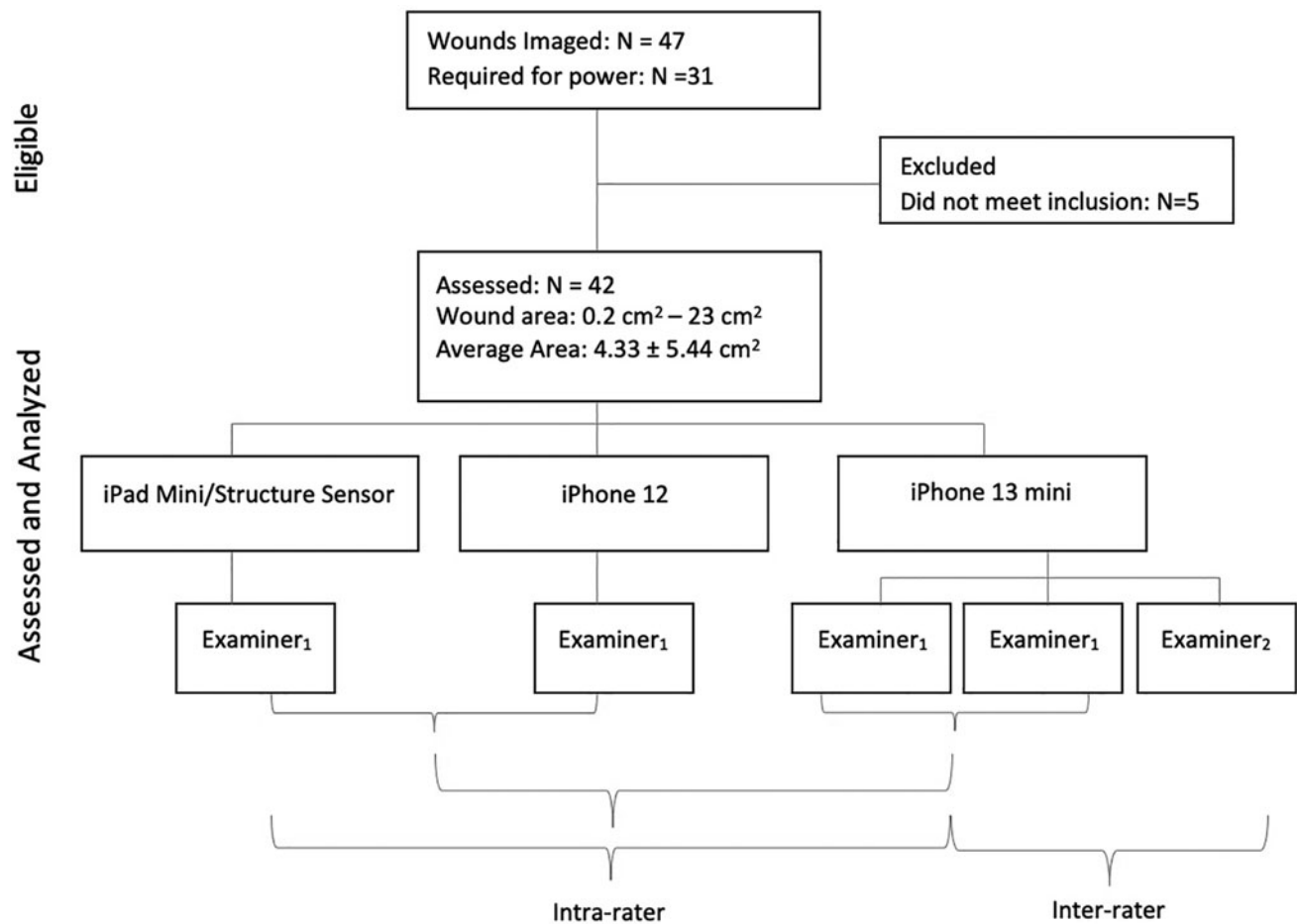


Figure 1. Summary graphic. Flow of wounds eligible for inclusion, assessed, and analyzed in the study.

of wound surface area. The assessments were analyzed using a two-way analysis of variance (ANOVA). Intraclass correlation coefficient (ICC) was computed for intrarater ($ICC_{1,1}$) and inter-rater ($ICC_{2,1}$) analysis using a two-way random effects model. Pearson's correlation was calculated between each imaging modality to assess the

strength of their linear relationship. A paired *t*-test was applied to assess the statistical difference between the measurement methods. Data were reviewed and tabulated in Excel 16.44 (Microsoft Corporation, Redmond, WA) and statistical analysis was performed using SAS 9.4 (SAS Institute, Inc., Cary, NC).



Figure 2. Modalities evaluated. iPhone 13 mini, 12, and iPad with Structure Sensor.



Figure 3. Automatic wound boundary delineation. Selected assessments showing automatic wound boundary delineation.

RESULTS

A total of 42 lesions from 30 patients were assessed in patients presenting with diabetic foot ulcers, venous ulcers, pressure injury, burn, autoimmune lesions, and iatrogenic wounds (Fig. 3). In two cases the same wound was imaged at multiple time points. The lesion surface area ranged from 0.2 to 23 cm² (average 4.33 ± 5.44 cm²). A high level of reliability was observed for repeat area measurements with iPhone 13 mini by the same examiner ($ICC_{1,1} = 0.997$). Similarly, a strong level of agreement was evidenced between examiners for wound surface area assessments ($ICC_{2,1} = 0.998$) (Table 1). ANOVA revealed that the differences in wound surface area for inter- and intraexaminer measurements ($p > 0.05$) were not statistically significant.

Table 1. Intra- and Inter-Rater Reliability

	ICC	p
Intrarater reliability	0.997	0.24
Inter-rater reliability	0.998	0.16

Intra- and inter-rater reliability is reported using an ICC.
ICC, intraclass correlation coefficient.

On validation of measurement methods, there was no statistically significant difference in area measurements between the iPhone 12 and iPad mini/Structure Sensor ($p = 0.78$), or between the iPhone 13 mini and iPhone 12 ($p = 0.22$). Comparison of area measurement between iPhone 13 mini and iPad mini with Structure Sensor, however, was at threshold ($p = 0.049$). Given the small effect size or standardized mean difference (Cohen's $d = 0.01$), this difference is unlikely to be discerned by an unaided eye. Moreover, the difference in mean of 1.39% is not clinically meaningful, especially considering that conventional methods of wound measurement (*e.g.*, ruler) have far larger variability¹⁷ (Table 2; Figs. 4–6).

DISCUSSION

In recent years, mobile devices have become ubiquitous, and there has been an increased utilization of smartphones in clinical practice. Significant benefits include point-of-care assessment and remote patient monitoring to better support clinical decision-making and observe outcomes. Both providers and patients are showing an increased reliance on digital apps; however, ensuring

Table 2. Agreement Between Devices

	13 vs. iPad/Sensor	12 vs. iPad/Sensor	13 vs. 12
Pearson correlation	0.999401477	0.998680006	0.998926671
$p(T \leq t)$ two-tail	0.048925747	0.781267975	0.222493848
Cohen's d	0.011125117		
Mean difference (%)	1.39		

Comparison of area measurements between iPhone 13 mini and 12, and 12 against iPad mini with Structure Sensor showed differences were not statistically significant. The difference between the iPhone 13 mini and the iPad mini with Structure Sensor was not considered clinically significant given the small effect size ($p=0.049$, Cohen's $d=0.01$).

quality and clinical validation across modalities is essential for confidence in results. Differences among devices may lead to variation in clinical measurements. Digital wound measurements, for instance, can be influenced by camera, position, and lighting conditions. This clinical study builds on prior work and aims to evaluate reliability across the new generation of iPhones.

Overall, a high level of agreement was observed for both measurement reliability and comparison of device configurations. The consistency and accuracy observed are likely due to the automatic

wound border delineation app feature. Increased variability may be seen when this feature is not in use or due to clinician differences in wound boundary judgment. Although the statistical difference between the iPhone 13 mini and iPad mini/Structure Sensor was at the threshold p -value, given the small effect size (Cohen's $d=0.01$), this difference is unlikely to be discerned or clinically significant. Conventional methods of wound measurement have greater variability.

An accurate assessment of wound area is critical for managing care and optimizing treatment plans. Although digital planimetry offers several benefits over the ruler method, it is still subject to variability. Improper placement of reference marker or sticker, angle between the camera and wound, and light conditions can influence the quality and accuracy of the assessment. In this study, we collected data quickly, accurately, and reproducibly without the need for a reference sticker using the inSight CR app on a standard iPhone. One study limitation to be addressed in future work is the collection of depth data for volumetric comparison.

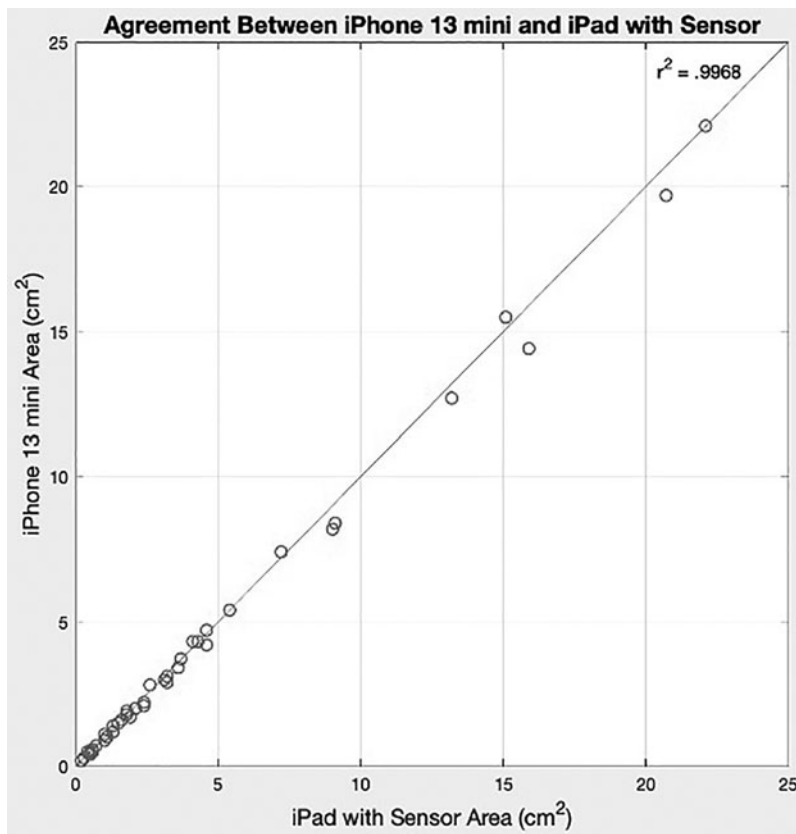


Figure 4. iPhone 13 mini versus iPad with Sensor. Comparison of examiner A's second set of images with the iPhone 13 mini versus iPad mini with Structure Sensor images. A high level of agreement was observed for area assessments (cm²).

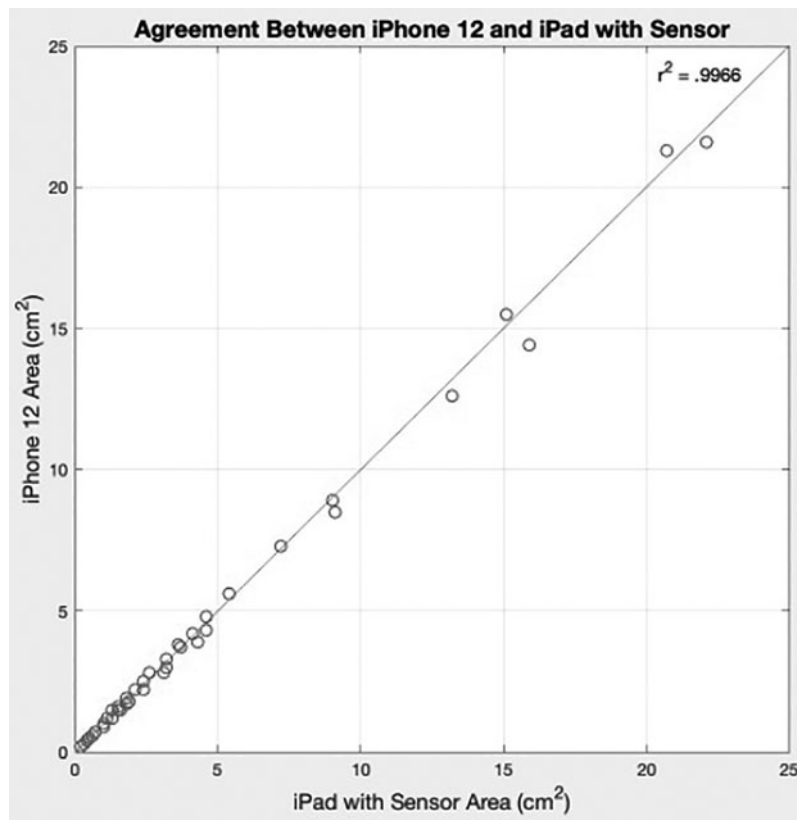


Figure 5. iPhone 12 versus iPad with Sensor. A high level of agreement was observed for area measurement (cm²) with the iPhone 12 and iPad mini with Structure Sensor.

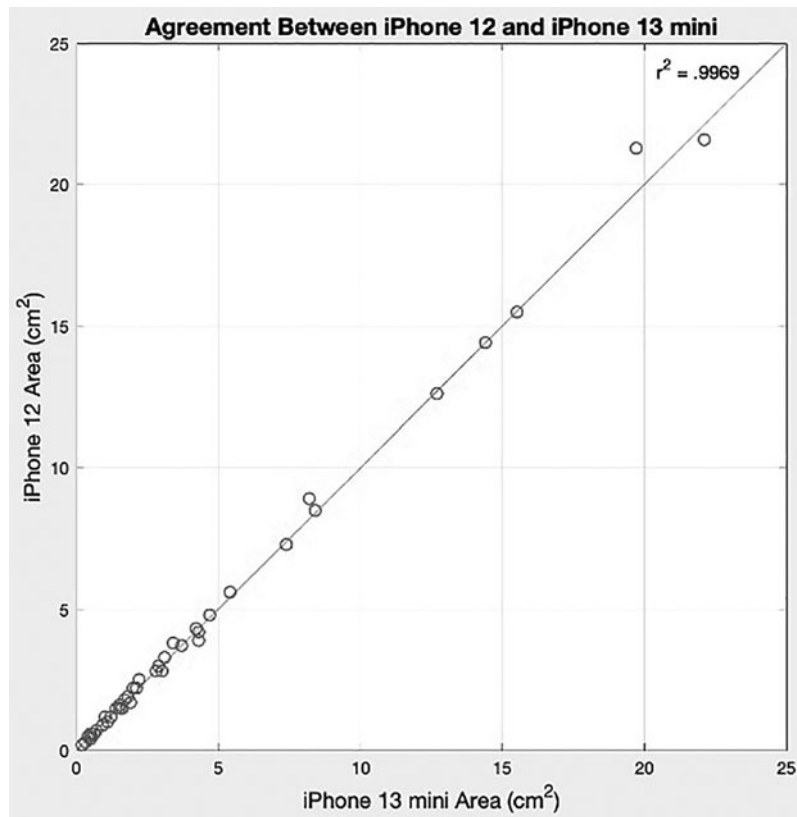


Figure 6. iPhone 12 versus iPhone 13 mini. Comparison of iPhone 12 images with examiner A's second set of images with the iPhone 13 mini. A high level of agreement was observed for area measurement (cm²).

Although there are no differences in the inSight CR app workflow between iPad mini/Structure Sensor and iPhone, various factors make the iPhone modality more convenient for the wound care clinician. First, a personal or work iPhone with the downloaded app can be used to image wounds, which is more common and less likely to require additional costs than an iPad or Structure Sensor. Furthermore, the Structure Sensor uses a rechargeable battery that lasts for a few hours of active sensing before needing to be recharged. There are no differences in the types of wounds able to be imaged. Wounds with difficult topography such as circumferential wounds were not included in this study.

Our results support the use of phone-based solutions for wound assessment. Smartphone-based wound measurement tools provide accurate and fast results, increased measurement capabilities for both providers in clinic and patients at home, and have a low barrier to widespread adoption. Based on the level of emphasis placed on iPhone quality assurance and control, it is reasonable to assume that these results are generalizable across this generation of devices. In addition, it is not unreasonable to assume a high level of concordance for future models with cameras and processing power that are at least comparable, if not superior to the iPhone 12 and 13.

INNOVATION

The findings provide evidence that the inSight CR app has a high level of reliability across devices and users. Differences between modalities were not a significant source of variation. Typically, current widespread wound measurement methods require tradeoffs between accuracy and efficiency. The proposed innovation demonstrated accurate wound measurement as compared with a previously validated measurement tool in seconds, with difference in iPhone 12 agreement not statistically significant ($p=0.78$) and minor iPhone 13 mini difference in agreement ($p=0.049$, Cohen's $d=0.01$) that is not clinically relevant. Wound results are also consistent between different examiners using the same device ($ICC_{2,1}=0.998$).

KEY FINDINGS

- There was no statistically significant difference in area measurements between the iPhone 12 and iPad mini/Structure Sensor ($p=0.78$), or between the iPhone 13 mini and iPhone 12 ($p=0.22$).
- A high level of reliability was observed for repeat area measurements with iPhone 13 mini by the same examiner ($ICC_{1,1}=0.997$).
- A strong level of agreement was evidenced between examiners for wound surface area assessments ($ICC_{2,1}=0.998$).

ACKNOWLEDGMENTS AND FUNDING SOURCES

The authors have no acknowledgements or funding sources to declare.

AUTHOR DISCLOSURE AND GHOSTWRITING

There are no financial or ghostwriting disclosures related to this article.

ABOUT THE AUTHORS

Mark Swerdlow, MS, is a 4th-year medical student at USC Keck pursuing plastic and reconstructive surgery. He has participated in 3 years of clinical wound-care and limb salvage research. Previously, he studied electrical engineering at Stanford where he completed a BS (bio-electronics/bio-imaging) and MS (signal processing).

Jessica Lo, BS, is a 2nd-year medical student at USC Keck. She graduated from Vanderbilt University with a double major in neuroscience and Spanish. She is involved in general and plastic and reconstructive surgery research and continues to explore her interests in various surgical fields.

David G. Armstrong, DPM, MD, PhD, is professor of surgery with tenure at the USC. He is founder and codirector of the Southwestern Academic Limb Salvage Alliance. Dr. Armstrong is the director of USC's National Science Foundation funded Center to Stream Healthcare in Place, which places him at the nexus of the merger of consumer electronics, wearables, and medical devices in an effort to maximize hospital-free and activity-rich days. He founded and cochairs the International Diabetic Foot Conference, the largest annual international symposium on the diabetic foot. He is also the founding president of the American Limb Preservation Society, dedicated to building interdisciplinary teams to eliminate preventable amputation.

REFERENCES

Abbreviations and Acronyms

3D = three-dimensional
ANOVA = analysis of variance
ICC = intraclass correlation coefficient
USC = University of Southern California